Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 2014345, 8 pages
https://doi.org/10.1155/2021/2014345

Research Article

Stacked Autoencoder Framework of False Data Injection Attack Detection in Smart Grid

Liang Chen,[1] **Songlin Gu**,[2] **Ying Wang**,[3] **Yang Yang**,[3] **and Yang Li**[4]

[1]*School of Automation, Nanjing University of Information Science and Technology, Nanjing 210044, China*
[2]*State Grid Economic and Technological Research Institute Co., Ltd, Beijing 102209, China*
[3]*State Grid Hebei Economic Research Institute, Shijiazhuang 050011, China*
[4]*School of Electrical Engineering, Northeast Electric Power University, Jilin 132012, China*

Correspondence should be addressed to Liang Chen; ch.lg@nuist.edu.cn

Received 20 May 2021; Accepted 25 June 2021; Published 3 July 2021

Academic Editor: Bo Yang

The advanced communication technology provides new monitoring and control strategies for smart grids. However, the application of information technology also increases the risk of malicious attacks. False data injection (FDI) is one kind of cyber attacks, which cannot be detected by bad data detection in state estimation. In this paper, a data-driven FDI attack detection framework of the smart grid with phasor measurement units (PMUs) is proposed. To enhance the detecting accuracy and efficiency, the multiple layer autoencoder algorithm is applied to abstract the hidden features of PMU measurements layer by layer in an unsupervised manner. Then, the features of the measurements and corresponding labels are taken as inputs to learn a softmax layer. Last, the autoencoder and softmax layer are stacked to form a FDI detection framework. The proposed method is applied on the IEEE 39-bus system, and the simulation results show that the FDI attacks can be detected with higher accuracy and computational efficiency compared with other artificial intelligence algorithms.

1. Introduction

Phasor measurement units (PMUs) can measure the voltage and current phasors directly with the help of global positioning system synchronization clock [1, 2]. Due to the ability of monitoring the transient dynamics of power systems, more and more PMUs have been installed in the smart grid. Meanwhile, the rapid developments of enhanced monitoring and information technology also facilitate the malicious cyber attacks [3]. The large-scale integration of renewable energy resources poses a challenge for the security of the system operation due to inherent uncertainties of renewables [4–6]. The cyber attacks on the power system monitoring and data acquisition systems are the main objectives for attackers to seriously threaten the power system operating safety. Attackers launch a cyber attack by sending a malicious information to the control center from measurements. One of the most important functions of a state estimator is bad data detection, by which some malicious attacks can be detected because the value of the objective function increases dramatically when attacks are launched. However, one kind of the serious cyber attacks that cannot be detected by bad data detection in state estimations is the false data injection (FDI) attack [7].

Up to now, lots of research works have been developed on different cyber attacks. Under the assumption that the network topology and parameters are known by the attackers, the FDI attack method is proposed in [8] for the first time. However, it is hard for the attacker to obtain the full acknowledgments of power systems. Aiming at this problem, in [9], a FDI attack method is given based on only partial knowledge of the system topology and a subset of meter measurements. To reduce attack costs and detection risks, the minimal set of meters that required to be compromised is taken as the objective function in [10]. In [11], the FDI attack is combined with other kind of cyber attacks, forming an enhanced FDI attack method. Once the FDI attack is launched in power systems, it is hard to be detected. To prevent the measurements being attacked, the meters should be protected. Lots of methods for minimizing the protection costs have been presented in [12, 13].

At the same time, the corresponding FDI attack detections are becoming a hot research topic. In [14], a reactance perturbation-based scheme is proposed to detect and identify originally covert FDI attacks on power system state estimation that enhances the security of state estimation without significantly increasing the operational cost in power systems. In [15], an online anomaly detection algorithm that utilizes load forecasts, generation schedules, and synchrophasor data to detect measurement anomalies is given. In [16], the feasibility and limitations of adopting the proactive false data detection approach to thwart FDI attacks on power grid state estimation are studied, and a framework to detect FDI attacks on power grid state estimation by using the proactive false data detection approach is proposed.

With the rapid developments of artificial intelligence technologies, the research works of data-driven technology-based detection methods are increasing dramatically. The principle component analysis is used to analyze the FDI attacks in the real-time environment [17], providing a more accurate and sensitive response than the previous FDI detection techniques. In [18], a supervised learning using labeled data called support vector machine-based FDI attacks detection method is proposed. The principal component analysis is used to reduce the dimension of the data to be processed, which leads to lower computation complexities. Use of deep learning for solving pattern classification problems is proven to be an effective way in engineering [19]. Under the FDI attack condition, spatial and temporal data correlations may deviate from those in normal operating conditions. Based on this characteristic, a discrete wavelet transform algorithm and deep neural networks' techniques are used to construct an intelligent system for AC FDI attack detection, which is proposed in [20]. In [21], the deep learning technique is applied to recognize the behavior features of FDI attacks with the historical measurement data and employ the captured features to detect the FDI attacks in real time. Although the deep learning is an effective method to detect the FDI attacks, some drawbacks, such as the heavy computation loads and bad generalization abilities with a huge amount of inputs, restrict the further applications. Autoencoders [22, 23] are one of the effective methods to cope with these problems, which can learn compressed features in an unsupervised manner, attracting more and more researchers' interests [24, 25]. However, the effectiveness of autoencoder decreases when the number of hidden units is more than the dimension of input data. To address this problem, sparse autoencoders, in which the sparsity is integrated into the autoencoder model to learn more efficient sparse features, have been developed [26]. In [27], a denoising autoencoder is used in wind turbine gearbox fault diagnosis, which can learn useful features from raw inputs by denoising. Due to the abilities of abstracting robust representations from noisy data, the denoising autoencoder is applied in many fields in recent years [27, 28]. In [29], autoencoders are used to reduce dimension and extract features from measurement datasets. Further, the autoencoders are integrated into an advanced generative adversarial network framework, which successfully detects anomalies under FDI attacks with a few labeled measurement data. However, the single-layer autoencoder cannot abstract entire representations of the original data. Aiming at this problem, a stacked autoencoder is proposed, which is made up of multiple autoencoders. The output of the first layer of the autoencoder is taken as the input of the second layer.

In this paper, a stacked autoencoder-based FDI attack detection framework in the smart grid is proposed. The main contributions are listed:

(1) A data-driven FDI attack detection framework is proposed. The topology errors and bad data are detected by state estimations. The hidden FDI attacks in measurements that cannot be identified by state estimation are detected by the intelligent algorithm.

(2) The stacked autoencoder is applied to detect the FDI attacks. Compared with other methods, the performances of the stacked autoencoder are better in the condition that the amounts of ordinary and attacks' samples differ widely.

(3) The proposed method is applied on the IEEE 39-bus testing system. The performances of the proposed method are better than the traditional deep learning methods, which are capable of practical applications.

The rest of this paper is organized as follows. Section 2 establishes the power system linear state estimation model. The bad data detection method is also given. In Section 3, the basic principle of FDI attacks is given. In Section 4, the stacked autoencoder-based FDI attack detection method is proposed. To evaluate the performance of the proposed FDI attack detection method, the case study is carried out under different conditions in Section 5. Finally, Section 6 concludes this paper.

2. Linear State Estimation of Power Systems

2.1. Linear State Estimation Model. With the rapid development of PMUs, it is possible to take the linear state estimation based on phasor measurements. The linear state estimation can be solved directly without iteration. As a result, the calculation burden of linear state estimation is lighter than nonlinear estimation. The measurements of linear state estimation include real and imaginary parts of bus voltages and currents phasors which can be measured directly. In the linear state estimation, the real and imaginary parts of bus voltages are taken as states that should be estimated. The relationships between branch current measurements and states are derived from the π equivalent of transmission lines, which are shown as follows:

$$\begin{cases} I_{ij,r} = \left(g_{ij} + g_{i0}\right)e_i - g_{ij}e_j - \left(b_{ij} + b_{i0}\right)f_i + b_{ij}f_j, \\ I_{ij,i} = \left(g_{ij} + g_{i0}\right)f_i - g_{ij}f_j + \left(b_{ij} + b_{i0}\right)e_i - b_{ij}e_j, \end{cases} \tag{1}$$

where $I_{ij,\mathrm{r}}$ and $I_{ij,\mathrm{i}}$ are the real and imaginary parts of the branch current phasors going from bus i to bus j, respectively, g_{ij} and b_{ij} are the conductance and susceptance of branch i-j, respectively, g_{i0} and b_{i0} are the conductance and susceptance of the shunt branch at bus i, respectively, and e_i and f_i are the real and imaginary parts of voltage phasor of bus i, respectively.

The matrix form of (1) is

$$\begin{bmatrix} I_{ij,r} \\ I_{ij,i} \\ I_{ji,r} \\ I_{ji,i} \end{bmatrix} = \begin{bmatrix} g_{ij} + g_{i0} & -b_{ij} - b_{i0} & -g_{ij} & b_{ij} \\ b_{ij} + b_{i0} & g_{ij} + g_{i0} & -b_{ij} & -g_{ij} \\ -g_{ji} & b_{ji} & g_{ji} + g_{j0} & -b_{ji} - b_{j0} \\ -b_{ji} & -g_{ji} & b_{ji} + b_{j0} & g_{ji} + g_{j0} \end{bmatrix} \begin{bmatrix} e_i \\ f_i \\ e_j \\ f_j \end{bmatrix}. \tag{2}$$

Equation (2) can be rewritten as

$$Z_B = H_B x, \tag{3}$$

where $z_B = [\ldots I_{ij,r}, I_{ij,i}, I_{ji,r}, I_{ij,i}, \ldots]^T$, $x = [\ldots e_i, f_i, e_j, f_j, \ldots]^T$, z_B is the vector of the branch current measurements, and x is the vector of states.

In addition to the branch current measurements, the injected currents and bus voltages can be measured by PMUs also. The measurement equation of linear state estimation is

$$\begin{bmatrix} z_U \\ z_B \\ z_{\mathrm{IN}} \end{bmatrix} = \begin{bmatrix} I_{2m\times 2n} \\ H_B \\ Y_M \end{bmatrix} x, \tag{4}$$

where z_U and z_{IN} are the phasor measurement vectors of bus voltages and injected currents, respectively, $I_{2m\times 2n}$ is the measurement matrix of bus voltages, m and n are the number of buses equipped with PMUs and the total bus number, respectively, and Y_M is the injected current measurement matrix.

Equation (4) can be rewritten as

$$z = Hx + v, \tag{5}$$

where z is the measurement vector, v is the measurement error, and v satisfies Gaussian distribution with zero mean and variance σ^2.

Equation (5) is linear, so the linear weighted least squares can be used to estimate the states. The objective function is to minimize the sum of weighted variances, which is shown as follows:

$$J(x) = [z - Hx]^T R^{-1} [z - Hx], \tag{6}$$

where J is the objective function, R is a diagonal matrix, the ith diagonal element of R is $1/\sigma_i^2$, and σ_i is the variance of ith measurement. The estimated states are

$$\hat{x} = \left[H^T R^{-1} H\right]^{-1} H^T R^{-1} z, \tag{7}$$

where $\hat{x}$ is the estimated states.

2.2. Bad Data Detection. Under the normal condition (no bad data in measurements), the sum of estimated measurement variance is under a given threshold ε; however, if the measurements experience bad data, the threshold ε would be exceeded. The sum of estimated measurement variance is given as

$$\begin{aligned} \hat{J} &= \hat{r}^T R^{-1} \hat{r}, \\ &= (z - \hat{z})^T R^{-1} (z - \hat{z}), \\ &= (z - H\hat{x})^T R^{-1} (z - H\hat{x}), \\ &= (z - Gz)^T R^{-1} (z - Gz), \\ &= z^T (I - G)^T R^{-1} (I - G) z, \end{aligned} \tag{8}$$

where $\hat{r}$ is the estimated measurement residual, $\hat{r} = z - \hat{z}$, $\hat{z}$ is the estimated measurement, $\hat{z} = H\hat{x}$, I is an identity matrix, and $G = H(H^{\mathrm{T}} R^{-1} H)^{-1} H^{\mathrm{T}} R^{-1}$.

The bad data can be detected by the following judgement:

$$\begin{cases} \hat{J} \le \varepsilon, & \text{no bad data,} \\ \hat{J} > \varepsilon, & \text{bad data exist.} \end{cases} \tag{9}$$

If the measurements experience bad data, the measurements would be removed one by one, and the states are estimated again until all bad data are removed.

3. False Data Injection Attacks

Aiming at the above bad data detection, FDI attack can construct an attack vector to the measurements that are able to bypass the bad data detection, but the estimated states deviate from the true values seriously. Assuming that the attackers can obtain the system typologies and parameters, the FDI attacks are formulated as follows:

$$z_a = z + a, \tag{10}$$

where z_a is the attacked measurement and a is the attack vector. If a is not artificially designed, the sum of estimated measurement variance would exceed the threshold, and the attack would be detected. As a result, the attacker must find out a proper vector a that will satisfy the following constrain:

$$\begin{aligned} \hat{r}_a - \hat{r} &= \left(z_a - H\hat{x}_c\right) - (z - H\hat{x}), \\ &= z_a - H(\hat{x} + c) - (z - H\hat{x}), \\ &= z + a - H\hat{x} - Hc - z + H\hat{x}, \\ &= a - Hc, \\ &= 0, \end{aligned} \tag{11}$$

where $\hat{r}_a$ is the estimated measurement residual under the bad data condition, $\hat{x}_c = \hat{x} + c$, $\hat{x}_c$ is the estimated states under attack condition, and c is the estimated deviation with attacked measurements. It can be seen from (11) that estimated measurement residual $\hat{r}_a$ under attack condition is equal to the residual $\hat{r}$ if the FDI attack vector a satisfies $a = Hc$. As a result, the FDI attack can bypass the bad data detection of (9). If the attacker obtains the overall structure and parameters, he can launch the attack by injecting malicious vectors to the measurements to change the estimating results as he wanted. This will cause serious consequences on power systems, while it cannot be detected.

The attacked measurements satisfy all constraints as the normal measurements, which can be presented as follows:

$$\begin{aligned} z_a &= z + a, \\ &= Hx + Hc, \\ &= H(x + c). \end{aligned} \tag{12}$$

Equation (12) shows that if the attacked measurement z_a satisfies constraints (5), the estimated states will deviate from actual values. This character leads to the hardness of detecting the FDI attacks using the traditional methods. In this paper, the stacked autoencoder is proposed to abstract the intrinsic features of the attacked measurements.

4. False Data Injection Attack Detection

4.1. Stacked Autoencoder. The autoencoder is a typical unsupervised learning neural network; the inputs of it are a set of unlabeled data. An autoencoder includes two parts: encoder and decoder. A reduced dimensional feature representation can be obtained by the encoder, which is taken as the inputs of decoders. The decoder tries to reconstruct the original input according to the reduced dimensional feature. The structure of the autoencoder is shown in Figure 1. z is the measurement vector, which is taken as inputs of the autoencoder. y is the reduced dimensional feature of z abstracted by the encoder, which is the decoder input. The output $\tilde{z}$ is the reconstruction of the original input z. The objective of the autoencoder is to try to copy its input to its output by two transformations:

$$\begin{aligned} y &= f\left(W_{1z} + b_1\right), \\ \tilde{z} &= g\left(W_{2y} + b_2\right), \end{aligned} \tag{13}$$

where f and g are the activation functions of the encoder and decoder, respectively, W_1 and W_2 are the weight matrixes, and b_1 and b_2 are the bias vectors.

W_1, W_2, b_1, and b_2 can be obtained by training the autoencoder using the unlabeled data z. It must be noted that the autoencoder can reconstruct different original inputs accordingly, which means that the feature representation y contains all information of the original input z in a lower dimensional form. As a result, the objective of the autoencoder is to minimize the gap between the output $\tilde{z}$ and input z. Thus, in the training process, the reconstruction loss function is

$$J_a\left(W_1, W_2, b_1, b_2,\right) = \arg\min \|z - \tilde{z}\|^2, \tag{14}$$

where J_a is the loss function of autoencoders.

In our FDI attack detection, once an autoencoder is trained, the output layer is useless. Only the hidden layer of the encoder is used to abstract the features of inputs. However, the application of a single encoder is limited. Aiming at this problem, the stacked autoencoder is proposed; the structure of it is shown in Figure 2. It can be seen that the outputs of one encoder are taken as the inputs of the next encoder. By this way, several encoders are stacked together to form a multilayer autoencoder. The features of original data are abstracted layer by layer. The stacked autoencoder is trained by the layer-wise unsupervised pretraining method. The encoder 1 is trained using the original data z by (14). The output of encoder 1 y_1 is taken as the input for training encoder 2. This process continues until the last encoder is trained. The output of each encoder is less than the former one. In the last, a softmax layer is trained by supervised learning using the output of the last encoder as input. The softmax layer function maps input scalars to a probability distribution; the values of it range from 0 to 1. The softmax layer is always used as the output layer for the classification problem. The probability function of the softmax layer is

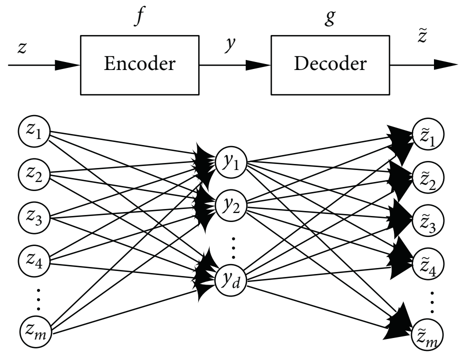


Figure 1: Basic structure of autoencoders.

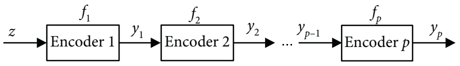


Figure 2: Structure of stacked autoencoders.

$$\phi(s) = \frac{e^{s_l}}{\sum_{c=1}^{C} e^{s_l}}, \quad l = 1, 2, \ldots, C, \tag{15}$$

where ϕ is the probability function of the softmax layer, s is the input of the softmax layer, s_l is the lth input element, and C is the total number of inputs. The sum of the softmax layer output elements is 1, and the value of each element represents the probability of the according classification.

4.2. Framework of False Data Injection Attack Detection. The flowchart of the proposed FDI attack detection is shown in Figure 3. After the measurement z_k is obtained, the linear state estimation should be taken first. Then, the value of the objective function is used to detect bad data. If the value exceeds the threshold, the bad data is deleted, and the state estimation is taken again, until all bad data are deleted. FDI attacks can bypass the bad data detection, so the proposed FDI attack detection is taken in the next step. If the attack is detected, the attacked measurements should be identified, which is not the research topic of this paper.

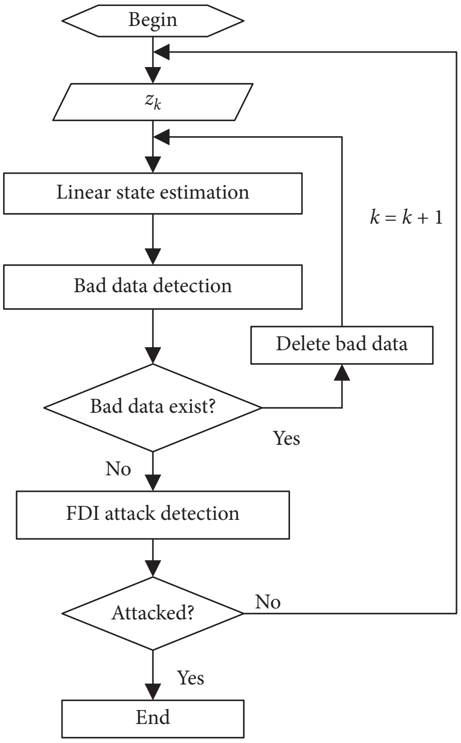


Figure 3: Flowchart of the FDI attack detection.

5. Case Studies

5.1. Descriptions of the Testing System and Data. To testify the validity of the proposed FDI attack detection method, the IEEE 39-bus testing system [16, 19] is used in this study. The voltage and current phasors can be measured by PMUs, which are taken as the inputs of the FDI attack detector. The power system states are obtained by power flow calculation using MATPOWER [30]. To simulate the practical operating condition, the generator and load powers are created by Monte Carlo simulations. The simulated values are true values, while the measured values are generated by adding specific distributed random numbers to the true values. The measurement errors of amplitudes and angles are 2% and 2°, respectively. Assume that the attacker chooses 5 states to be attacked, and the estimated deviation c ranges from −2 to 2. The attacked value $a = Hc$ is added to measurement z to form z_a. In practice, the attacked measurements are far less than the normal measurements. In this simulation, the training set includes 5000 normal measurement samples and 500 attacked samples; the testing set includes 3000 normal samples and 300 attacked samples.

In this study, two encoders and a softmax layer are stacked to form the stacked autoencoder-based FDI attack detection framework. The overall structure as well as the input and output numbers of the stacked encoders are shown in Figure 4.

5.2. The Performances of the Method. To evaluate the performance of the detection method, the confusion matrix is used to analyze the detection results quantitatively, which are defined in Figure 5. The true positives (TP) means that actual attacks are correctly classified as attacks; the true negatives (TN) means that actual normal measurements are correctly classified as no attack; the false positives (FP) means that actual normal measurements are incorrectly classified as attacks; the false negatives (FN) means that actual attacks are incorrectly classified as no attacks. The following three indexes are used to evaluate the ability of the proposed method, which are defined as

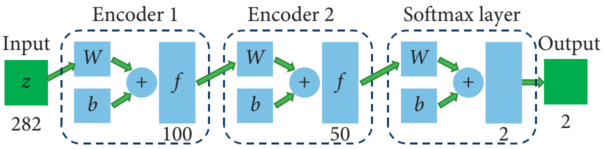


Figure 4: Structure of the stacked encoder.

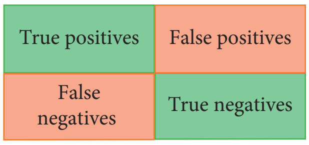
$$\begin{aligned} \text{Acc} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \\ \text{Pre} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\ \text{Rec} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \end{aligned} \tag{16}$$

where Acc, Pre, and Rec are the accuracy, precision, and recall, respectively, Acc represents the overall performances of the method, Rec evaluates performances of the attack detection, and Pre evaluates the probability that the normal measurements are not detected as attacks.

The confusion matrix of the detection results is shown in Figure 6. It can be seen that the 300 attacks are detected out; the others are detected as normal measurements. The index values of Acc, Pre, and Rec are 100%, 100%, and 100%, respectively.

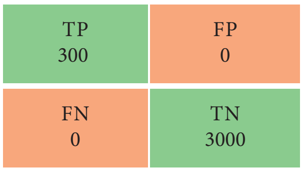
5.3. Comparison with Other Methods. Three other detection methods, i.e., multilayer perceptron (MLP), support vector machines (SVM), and deep neural network (DNN), are applied in the simulation. The neuron number in the hidden layer of MLP is 15. If the output of MLP is smaller than 0.5, the classification is no attack; otherwise, the classification is being attacked. For the DNN, the number of hidden layers is 4, and the unit number of each hidden layer is 150. The confusion matrixes and the methods are shown in Figure 7. It shows that the TN numbers of the three methods are 3000, meaning that all normal measurements are correctly detected. However, the 300 attacks are not detected accurately; the detection performance of which can be evaluated by the index of Rec shown in Table 1. Among the three methods, the performance of the DNN method is better than the other two methods. However, it is still worse than the proposed detection method.

5.4. Sensitivity Analysis. In this section, the influences of the following factors to the detection performances will be studied:



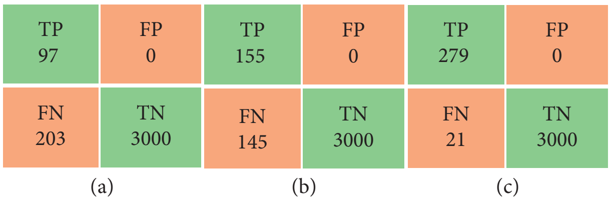
True positives	False positives
False negatives	True negatives

Figure 5: Confusion matrix.



TP 300	FP 0
FN 0	TN 3000

Figure 6: Confusion matrix of detection results.



TP 97	FP 0
FN 203	TN 3000

(a)

TP 155	FP 0
FN 145	TN 3000

(b)

TP 279	FP 0
FN 21	TN 3000

(c)

Figure 7: Confusion matrixes of three different methods. (a) MLP. (b) SVM. (c) DNN.

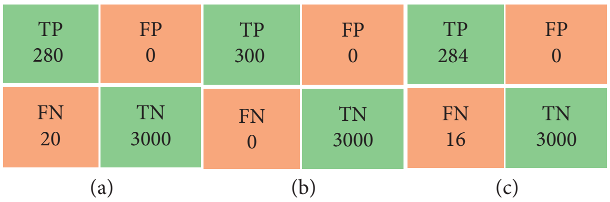
Table 1: The values of indexes of three methods.

	Acc (%)	Pre (%)	Rec (%)
MLP	93.8	100.0	32.3
SVM	95.6	100.0	51.7
DNN	99.4	100.0	93.0
Stacked encoder	100	100	100

(1) The number of neurons: the influence of neuron numbers of the encoders is studied in this section. The neuron numbers of encoder 1 and encoder 2 are set in the following cases:

Case 1: 50 and 10
Case 2: 80 and 40
Case 3: 20 and 200

The confusion matrixes of detection results are shown in Figure 8. It shows that 20 attacks are not detected in Case 1, meaning that the performance of the proposed method decreases if the neuron is less. In Case 3, 16 attacks are not detected. The reason is that the neuron number of encoder 1 is 20, which cannot abstract the full features in the measurements, although the neuron number of encoder 2 is 200.



TP 280	FP 0
FN 20	TN 3000

(a)

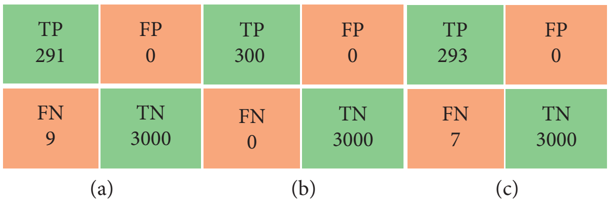
TP 300	FP 0
FN 0	TN 3000

(b)

TP 284	FP 0
FN 16	TN 3000

(c)

Figure 8: Confusion matrixes of different neurons. (a) Case 1. (b) Case 2. (c) Case 3.



TP 291	FP 0
FN 9	TN 3000

(a)

TP 300	FP 0
FN 0	TN 3000

(b)

TP 293	FP 0
FN 7	TN 3000

(c)

Figure 9: Confusion matrixes of different encoder numbers. (a) Case 1. (b) Case 2. (c) Case 3.

(2) The number of encoders: the influence of the encoder number stacked in the detection algorithm is studied. The following 3 cases are considered:

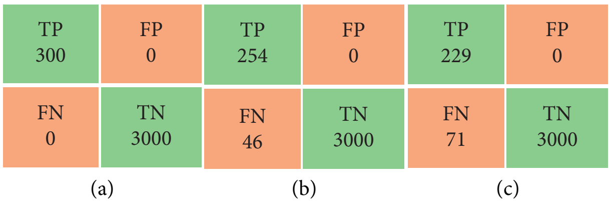
Case 1: 1 encoder; 100 neurons
Case 2: 3 encoders; 200, 100, and 50 neurons for each encoder
Case 3: 3 encoders; 50, 20, and 10 neurons for each encoder

The confusion matrixes of detection results are shown in Figure 9. It can be seen that 9 attacks are not detected in Case 1 because there is only one encoder, and the features cannot be abstracted fully. Although there are 3 encoders in Case 3, 7 attacks are not detected because the neurons of each encoder are less.

(3) Attack proportions of the training set: in practice, the attacked samples are much less than the normal samples. The influence of attack proportions in the training set is studied also. The detection framework of Figure 4 is applied, and the testing samples include 3000 normal measurements and 300 attacks. The following training sets are considered:

Case 1: 7000 normal samples; 500 attacks
Case 2: 9000 normal samples; 500 attacks
Case 3: 9500 normal samples; 200 attacks

The confusion matrixes are shown in Figure 10. It shows that, with the decreasing proportion of attack samples, more attacks cannot be detected. The proposed method is sensitive to the proportion of attacks in the training set. The reason is that the features of FDI attacks are hard to be abstracted by the encoder when the attack proportion is low.



TP 300	FP 0
FN 0	TN 3000

(a)

TP 254	FP 0
FN 46	TN 3000

(b)

TP 229	FP 0
FN 71	TN 3000

(c)

Figure 10: Confusion matrixes of different attack proportions. (a) Case 1. (b) Case 2. (c) Case 3.

6. Conclusion

In this paper, a stacked autoencoder-based FDI attack detection framework is proposed, and it is applied on the IEEE 39-bus testing system under different conditions. The confusion matrix and 3 indexes are used to evaluate the performances of the detection methods. The simulation results show that the neuron numbers of encoders influence the detection performance. If the neurons are less, the features cannot be abstracted fully, resulting in the low Rec values. The encoder number is another aspect influencing the detection performances. If the encoders are less, some attacks cannot be detected. It should be noted that if the neurons are less, the detection performances still decrease even when many encoders are stacked. The proposed detection method is sensitive to the attack sample proportion in the training set. If too few attacks are in the training sets, the features of FDI attacks cannot be abstracted fully, and the detection performance is decreased.

The FDI attack detection based on stacked autoencoders can be carried out in the following areas: the method of determining the optimal number of encoders and neurons, denoising function of the detectors, robustness to the wrong labeled samples, and detection with unbalanced data. Another interesting topic is to extend this work for detecting cyber attacks in integrated energy systems [31–36].

Data Availability

The IEEE 39-bus system data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partly supported by the Natural Science Foundation of Jilin Province, China, under Grant no. 2020122349JC.

References

[1] A. G. Phadke, "Synchronized phasor measurements in power systems," *IEEE Computer Applications in Power1993*, vol. 6, no. 2, pp. 10–15, 1993.

[2] Y. Li and Z. Yang, "Application of EOS-ELM with binary Jaya-based feature selection to real-time transient stability assessment using PMU data," *IEEE Access*, vol. 5, pp. 23092–23101, 2017.

[3] Y. Xiang, L. Wang, and N. Liu, "Coordinated attacks on electric power systems in a cyber-physical environment," *Electric Power Systems Research*, vol. 149, pp. 156–168, 2017.

[4] Y. Li, Z. Yang, G. Li, D. Zhao, and W. Tian, "Optimal scheduling of an isolated microgrid with battery storage considering load and renewable generation uncertainties," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 2, pp. 1565–1575, 2019.

[5] J. Li, B. Lu, Z. Wang, and M. Zhu, "Bi-level optimal planning model for energy storage systems in a virtual power plant," *Renewable Energy*, vol. 165, pp. 77–95, 2021.

[6] Y. Li and K. Li, "Incorporating demand response of electric vehicles in scheduling of isolated microgrids with renewables using a bi-level programming approach," *IEEE Access*, vol. 7, pp. 116256–116266, 2019.

[7] Y. Li, Z. Li, and L. Chen, "Dynamic state estimation of generators under cyber attacks," *IEEE Access*, vol. 7, pp. 125253–125267, 2019.

[8] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security*, vol. 14, no. 1, pp. 1–33, 2011.

[9] M. A. Rahman and H. Mohsenian-Rad, "False data injection attacks with incomplete information against smart power grids," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, pp. 3153–3158, Anaheim, CA, USA, December 2012.

[10] K. C. Sou, H. Sandberg, and K. H. Johansson, "Electric power network security analysis via minimum cut relaxation," in *Proceedings of the 2011 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pp. 4054–4059, Orlando, FL, USA, December 2011.

[11] M. N. Kurt, Y. Yilmaz, and X. Wang, "Real-time detection of hybrid and stealthy cyber-attacks in smart grid," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 2, pp. 498–513, 2019.

[12] J. Chen and A. Abur, "Placement of PMUs to enable bad data detection in state estimation," *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1608–1615, 2006.

[13] T. T. Kim and H. V. Poor, "Strategic protection against data injection attacks on power grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 326–333, 2011.

[14] C. Liu, J. Wu, C. Long, and D. Kundur, "Reactance perturbation for detecting and identifying FDI attacks in power system state estimation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 4, pp. 763–776, 2018.

[15] A. Ashok, M. Govindarasu, and V. Ajjarapu, "Online detection of stealthy false data injection attacks in power system state estimation," *IEEE Transactions on Smart Grid*, vol. 9, no. 3, pp. 1636–1646, 2019.

[16] B. Li, G. Xiao, R. Lu, R. Deng, and H. Bao, "On feasibility and limitations of detecting false data injection attacks on power grid state estimation using D-FACTS devices," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 854–864, 2020.

[17] A. S. Musleh, M. Debouza, H. M. Khalid, and A. Al-Durra, "Detection of false data injection attacks in smart grids: a real-time principle component analysis," in *Proceedings of the IECON 2019—45th Annual Conference of the IEEE Industrial Electronics Society*, pp. 2958–2963, Lisbon, Portugal, October 2019.

[18] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, “Detecting stealthy false data injection using machine learning in smart grid,” *IEEE Systems Journal*, vol. 11, no. 3, pp. 1644–1652, 2017.

[19] M. Zhang, J. Li, Y. Li, and R. Xu, “Deep learning for short-term voltage stability assessment of power systems,” *IEEE Access*, vol. 9, pp. 29711–29718, 2021.

[20] J. J. Q. Yu, Y. Hou, and V. O. K. Li, “Online false data injection attack detection with wavelet transform and deep neural networks,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3271–3280, 2018.

[21] Y. He, G. J. Mendis, and J. Wei, “Real-time detection of false data injection attacks in smart grid: a deep learning-based intelligent mechanism,” *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2505–2516, 2017.

[22] Y. Lin and J. Wang, “Probabilistic deep autoencoder for power system measurement outlier detection and reconstruction,” *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1796–1798, 2020.

[23] A. Majumdar, “Blind denoising autoencoder,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 1, pp. 312–317, 2019.

[24] X. Han, Y. Zhong, and L. Zhang, “Spatial-spectral unsupervised convolutional sparse auto-encoder classifier for hyperspectral imagery,” *Photogrammetric Engineering and Remote Sensing*, vol. 83, no. 3, pp. 195–206, 2017.

[25] G. Abdi, F. Samadzadegan, and P. Reinartz, “Spectral-spatial feature learning for hyperspectral imagery classification using deep stackedsparse autoencoder,” *Journal of Applied Remote Sensing*, vol. 11, no. 4, pp. 1–15, 2017.

[26] J. Deng, Z. Zhang, E. Marchi, and B. W. Schuller, “Sparse autoencoder based feature transfer learning for speech emotion recognition,” in *Proceedings of Humaine Association Conference on Affective Computing and Intelligent Interaction*, pp. 511–516, Geneva, Switzerland, September 2013.

[27] G. Jiang, H. He, P. Xie, and Y. Tang, “Stacked multilevel-denoising autoencoders: a new representation learning approach for wind turbine gearbox fault diagnosis,” *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 9, pp. 2391–2402, 2017.

[28] J. Yu, “Enhanced stacked denoising autoencoder-based feature learning for recognition of wafer map defects,” *IEEE Transactions on Semiconductor Manufacturing*, vol. 32, no. 4, pp. 613–624, 2019.

[29] Y. Zhang, J. Wang, and B. Chen, “Detecting false data injection attacks in smart grids: a semi-supervised deep learning approach,” *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 623–634, 2021.

[30] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, “MATPOWER: steady-state operations, planning, and analysis tools for power systems research and education,” *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

[31] Y. Li, C. Wang, G. Li, and C. Chen, “Optimal scheduling of integrated demand response-enabled integrated energy systems with uncertain renewable generations: a Stackelberg game approach,” *Energy Conversion and Management*, vol. 235, Article ID 113996, 2021.

[32] Y. Liu, Y. Li, H. B. Gooi et al., “Distributed robust energy management of a multimicrogrid system in the real-time energy market,” *IEEE Transactions on Sustainable Energy*, vol. 10, no. 1, pp. 396–406, 2019.

[33] Y. Li, C. Wang, G. Li, J. Wang, D. Zhao, and C. Chen, “Improving operational flexibility of integrated energy system with uncertain renewable generations considering thermal inertia of buildings,” *Energy Conversion and Management*, vol. 207, Article ID 112526, 2020.

[34] Y. Li, T. Zhao, P. Wang et al., “Optimal operation of multimicrogrids via cooperative energy and reserve scheduling,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3459–3468, 2018.

[35] Y. Li, Z. Yang, D. Zhao, H. Lei, B. Cui, and S. Li, “Incorporating energy storage and user experience in isolated microgrid dispatch using a multi-objective model,” *IET Renewable Power Generation*, vol. 13, no. 6, pp. 973–981, 2019.

[36] Y. Li, Z. Han, D. Yang, and G. Li, “Coordinating flexible demand response and renewable uncertainties for scheduling of community integrated energy systems with an electric vehicle charging station: a bi-level approach,” *IEEE Transactions on Sustainable Energy*, 2021.